B Coding Manual

This coding manual is designed for analyzing liquid variation in Spanish, specifically the variable production of /ɾ/, /r/, and /l/, in words from interviews in the **Boston Spanish Corpus (BSC)**. Originally developed for Lee-Ann Vidal-Covas' dissertation within the Spanish in Boston Project, the manual provides detailed guidelines for data collection and analysis of the **BSC** interviews.

Data Preparation

To ensure systematic and reproducible analysis, follow the process detailed below. The study uses audio recordings of sociolinguistic interviews, which are transcribed orthographically in *Praat*. These transcriptions are paired with audio to identify and annotate lexical items containing liquid segments.

1. Identifying and Segmenting Liquid Tokens

   Lexical items with liquid tokens are identified and manually segmented in Praat text grids. Each host word containing a liquid segment should be segmented on a designated tier. A minimum of 250 tokens must be collected per speaker. While identifying tokens for taps and laterals is relatively straightforward, trill tokens require special attention due to their lower frequency.

2. Annotating Liquid Segments

After segmenting the host words, create an additional text grid tier for annotating the liquid segments within each host word. The annotation process follows the segmentation protocol detailed in the Segmenting liquids Section.

3. Adding Predictor Variable Tiers

   Additional tiers, as described in the Tier Descriptions section, are added to systematically code predictor variables for all tokens. This structured annotation ensures the dataset's integrity and supports robust acoustic and variationist analysis in subsequent stages.

Textgrid Tiers

Below is a list and description of the interval tiers to be added to the **Praat** (Boersma & Weenink 2020) TextGrid for data collection. These tiers represent the dependent and independent variables used to study liquid variation. While most variables are coded manually during TextGrid preparation, some independent variables are automatically extracted into a CSV file using a *Praat Script* after the coding process is complete. Detailed instructions for identifying sites of variation, segmenting liquid sounds, and applying the coding schema are provided in the sections that follow.

List of Tiers

The following tiers are added to the TextGrid during annotation:

- Interviewer
- Orthographic

- Host Word
- Number of Tokens per Host Word
- Phonological Form
- Surface Form
- Number of Syllables
- Syllable Stress
- Syllable Type
- Position in Syllable
- Position in Word
- Word Class
- Previous Sound
- Following Sound

Added with Praat extraction script After the TextGrids are completed, the following measurements are generated automatically using the *Praat Script* `Liquids_Data_Extraction.praat`, which compiles data from all TextGrids into a clean CSV file:

- Duration of Host-Word
- Acoustic Analysis (Formants at 11 points)

Added in R Following the data extraction, the following variables are added during post-processing in **R**. The script `Create_liquids_master_df.Rmd` automates this step by calculating and incorporating these variables:

- Lexical Frequency

- SPEECH RATE

$$\text{SPEECH RATE (IN MS PER SYLLABLE)} = \frac{\text{DUR(MS) OF HOST-WORD}}{\#\text{ OF SYLLABLES}}$$

Tier Descriptions

This section provides detailed descriptions of each tier included in the TextGrid annotation process. These tiers are critical for ensuring consistent and accurate data coding and represent the dependent and independent variables used in the analysis. Each tier description outlines its purpose, how it should be coded, and any relevant considerations or notes. By following these guidelines, the annotation process becomes more systematic, reducing the likelihood of errors and enabling efficient post-processing.

Orthographic Orthographic transcription of the consultant's speech, which is used when segmenting the host words that include liquids. This tier will not need to be added, as the conversations have been transcribed by the time the liquids are to be extracted. It will also not appear in the data frame.

Interviewer Orthographic transcription of the interviewer's speech. This tier will also be completed previous to liquid data collection and will not appear in the data frame.

Host Word Interval of speech which includes the lexical item that hosts the liquid tokens. Interval boundaries are essential as the duration of *Host Word* will be used to compute the variable *Speech Rate*.

Number of Tokens per host word Total number of liquid tokens in the host word. Examples are provided Table B.1.

Table B.1: Tokens per Word Examples

Host Word	Translation	# of tokens
puerto	port	1
latinoamérica	Latin America	2
otorrinolaringologo	otolaryngologist	4

Phonological Form This variable codes the presumed phonological identity of the token as one of three categories: /ɾ/, /r/, or /l/. We use the orthographic spelling of the host word, its presumed phonological form, and an expanded version of Hualde's (2014: 183) guide to Spanish rhotic distribution to differentiate between the two rhotics: /ɾ/ and /r/. The expanded version introduces specific codes (*e.g.*, a1 – d3) to facilitate working in

Table B.2: Hualde's (2014) Spanish Rhotic Distribution Expanded

Environment	Code	Context	Example	Translation
Contrastive *(word-dependent)*				
V__V	a1 a2	Intervocalic & Word-internal	/ka.ɾo/ vs. /ka.ro/	expensive vs. car
/r/				
#__	b1	Word-initial (or root-initial)	/ro.ka/	rock
C.__	b2	After heterosyllabic consonant	/al.re.de.dor/	around
/ɾ/				
#C__V	c1	After tautosyllabic consonant; syllable onset; word-initial	/bɾo.ma/	joke
C.C__	c2	After tautosyllabic consonant; syllable onset; word-internal; consonant in previous syllable coda	/en.tɾe/	between
V.C__	c3	After tautosyllabic consonant; syllable onset; word-internal; vowel in previous syllable coda	/a.bɾu.mar/	to overwhelm
V__#V	c4	Word final followed by vowel	/seɾ#a.mi.gos/	to be friends
Variable *(usually /ɾ/)*				
V__C	d1	After vowel & Before consonant	/paɾ.te/	part
V__#C	d2	Word final followed by consonant	/seɾ#po.e.ta/	to be a poet
V__##	d3	Word final followed by pause	/seɾ/	to be

[1] # represents word boundary
[2] ## indicates a pause or phonological phrase boundary
[3] . represents a syllable boundary

Praat, streamlining the process of identifying and classifying instances of rhotic variation.

Additionally, new contexts (*i.e.*, a1 – a2 & c1 – c3) and corresponding examples have been incorporated. These additions enable simultaneous coding of multiple independent variables when coding for Phonological Form, thereby saving time and enhancing accuracy. See Table B.2 for the full coding schema.

Important Notes for Coding Phonological Form

- Please use the codes in the **Code** column when coding in Praat.
- When coding for the lateral /l/:
  - Add an "l" at the beginning of codes b–d
  - For the **Environment** V__V (codes a1 & a2), just write "la" and exclude the number.

Surface Form The *perceived* segmental identity of the liquid token. For this tier, the researcher will code what they hear, even if different from the reading in the spectrogram and waveform. This variable has the values:

[ɾ], [r], [r̝], [ɹ], [ɽ], [l], [ɬ], [x], [v], [ð], [i], [h], [ʔ], [∅]

Number of Syllables The syllable count of each host word need to be coded to calculate speech rate; see Table B.3 for examples.

Table B.3: Number of Syllables in Word

Host Word	Translation	# of Syllables
las	the (fem. plural)	1
puer.to	port	2
la.ti.no.a.mé.ri.ca	Latin America	7
o.to.rri.no.la.rin.gó.lo.go	otolaryngologist	9

Syllable Stress Whether the syllable to which the token belongs bears primary stress. The values are *stressed* or *unstressed*. Examples are provided in Table B.4.

Table B.4: Syllable Stress Example

HOST WORD	SYLLABIFICATION	STRESS
diferente	di.fe.ˈ**ren**.te	stressed
escuela	es.ˈcue.**la**	unstressed
gustaría	gus.ta.ˈ**rí**.a	stressed
otras	ˈo.**tras**	unstressed
problema	pro.ˈ**ble**.ma	stressed
quieres	ˈquie.**res**	unstressed

- Primary stress is marked with a stress diacritic (ˈ) placed before the stressed syllable.
- Syllables containing the liquid token are **bolded**.
- Monosyllabic words should be coded as *stressed*.
- Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Syllable Type Whether the token belongs to an *open* or *closed* syllable. Open syllables end in a vowel sound, whereas closed syllables end with a consonant sound. Examples are provided in Table B.5.

1. open
2. closed

Note: Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Table B.5: Syllable Type Examples

HOST WORD	TRANSLATION	FIRST SYLLABLE	SYLLABLE TYPE	SECOND SYLLABLE	SYLLABLE TYPE
clase	class	cla	open	se	open
parte	part	par	closed	te	open

Position in Syllable Position of the liquid token within the syllable. Examples are listed in Table B.6. The values are:

1. onset
2. coda

Note: Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Position in Word Position of the liquid token within the within the host word. Refer to Table B.7 for examples. The values are:

Table B.6: Position in Syllable Examples

Host Word	Translation	First Syllable	Position in Syllable	Second Syllable	Position in Syllable
real	royal / real	re	onset	al	coda
partir	to leave / to split	par	coda	tir	coda

1. initial
2. internal
3. final

Note: Please use the numbers that belong to each value when coding in Praat. This will avoid misspellings and cut down on coding time.

Table B.7: Position in Word Examples

Host Word	Translation	First Syllable	Position in Word	Second Syllable	Position in Word
real	royal / real	re	initial	al	final
partir	to leave / to split	par	internal	tir	final

Word Class Word class to which the host word belongs. The same tier will also code for specific features in some classes to allow for post-hoc research, as seen in Table B.8.

Previous Sound The identity of the phonetic segment immediately preceding the liquid token.

- [a], [b], [β], [d], [ð], [e], [f], [g], [ɣ], [h], [i], [j], [k], [l], [m], [n], [ŋ], [o], [p], [k], [r], [ɾ], [ʀ], [ɽ], [ɹ], [s], [t], [u], [v], [x], [z]
- "CT" for when the sounds are unrecognizable
- "Pause" for when there is no sound (i.e. there is a pause)
- These will be recoded to create new variables in R:
  - if *vowel*: Vowel Height, Vowel Displacement
  - if *consonant*, Place & Manner of Articulation

Table B.8: Word Class Coding Schema

CLASS	CODE	FUNCTION	TYPE	PRONOUN+	EXAMPLES
Noun	1		proper (p) / common (c)		Puerto Rico, carro
Verb	2		infinitive (i) conjugated(c) past participle (t) gerund (g)	P	celebramos, celebrar, celebrarle Review of Spanish Participles
Determiner	3	modify nouns			el la los las un una unos unas todos algunos estos aquellos su sus numerals (primero, segundo)
Adjective	4	describe nouns			grande, largo, corto, hermoso, azul
Adverb	5	describe verbs			lentamente, rapidamente
Pronoun	6				yo, tú, él, ella
Preposition	7	explain relationships b/w nouns			a, ante, bajo, cabe, con, contra, de, desde, en, entre, hacia, hasta, para, por, según, so, sobre, tras
Conjunction	8	connect nouns, verbs, phrases, & sentences			y, o, pero, sin embargo, por lo tanto, asi que, aunque, porque, para que
Other	9	catch-all			Use this for anything that doesn't fall into any of the previous categories

Note: PRONOUN+ will be coded for in the case of verbs with a suffixed pronoun, i.e. the verb *celebramos* would be coded as 2c while *celebrarle* would be coded as 2iP

Following Sound The identity of the phonetic segment immediately following the liquid token. (This variable has the same possible values as the variable PREVIOUS SOUND, above.)

Identifying the Tokens

Inclusions

When identifying tokens, it is important to include all liquid segments present in a word. The extraction script is designed to accommodate multiple tokens from a single word, ensuring that no relevant data is overlooked. This comprehensive approach allows for the systematic analysis of liquid variation within and across words.

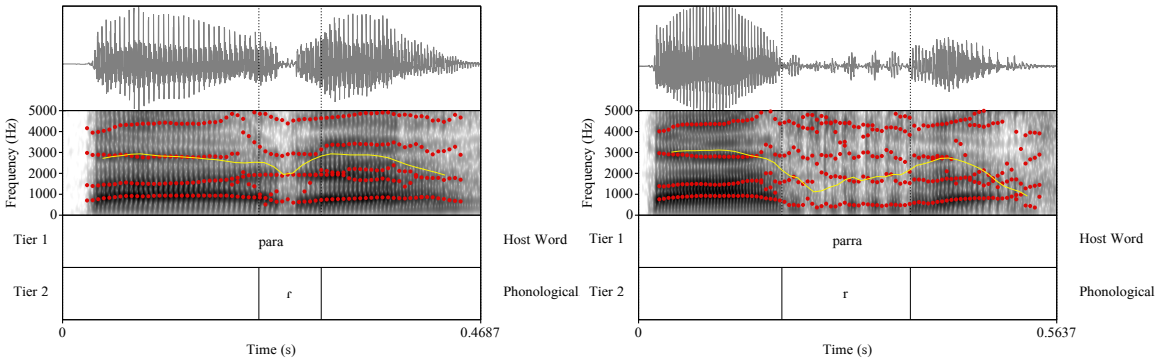
Segmenting liquids

Following the data preparation process, the segmentation of host words and individual liquid tokens is to be carried out in Praat. Each liquid token should be carefully segmented by closely examining both the waveform and the spectrogram to distinguish it from the surrounding sounds. In the segmentation process, two key objectives guide this protocol: (1) distinguishing the liquid from the surrounding sounds, and (2) classifying its identity based on its acoustic properties. These steps are essential for accurately analyzing the tokens, particularly in cases where the realization of liquids diverges from canonical expectations. The *Surface Form* variable, central to this study, captures such variation by providing a structured framework for classifying these sounds, even when they exhibit atypical acoustic characteristics.

The waveform, which represents frequency and amplitude over time, was analyzed to assess periodicity. Periodic speech sounds exhibit complex repeating waves, indicating regular vocal fold vibration, while aperiodic sounds display random, non-repetitive patterns. It should be noted that periodic and aperiodic components can co-occur within the same sound, as is often the case in speech. Spectrograms provide more detailed visual cues for identifying boundaries between sounds by displaying formant structure. For all liquids, changes in formant frequency—such as shifts in the first (F1), second (F2), or third (F3) formants—provide critical information for distinguishing between the preceding and following sounds. These changes often reflect the articulatory transitions between the liquid and adjacent sounds, such as the movement of the tongue or the opening and closing of the vocal tract, which can cause abrupt or gradual shifts in the formant structure. Addition-

ally, intensity changes, such as a dip in amplitude when transitioning from a vowel to a liquid, are key cues for segmentation. The next section shows some examples to illustrate the segmentation and coding procedure for both rhotics and laterals.

Rhotics Consider the two rhotics in Figure B·1. Figure B·1a includes a flap in the word [para], while Figure B·1b includes a trill in the word [para].



(a) Pa[ɾ]a example (b) Pa[r]a example

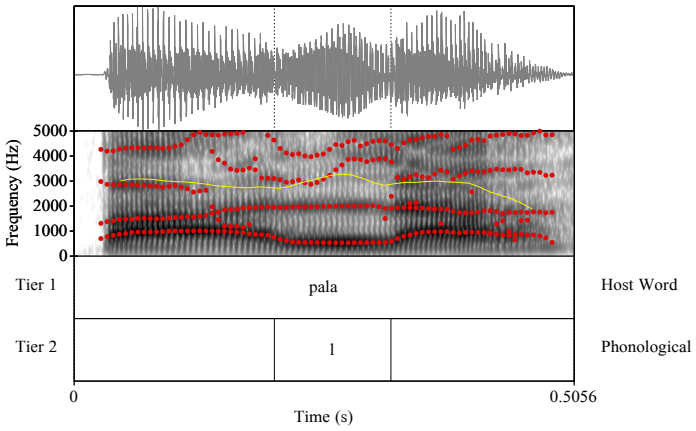
Figure B·1: Waveform, spectrogram, and annotation of pa[ɾ]a and pa[r]a

- Some possible cues to segment rhotics are:
  - There is a-periodicity in the waveform.
  - A steep fall of the third formant is a defining acoustic signature of rhotic sounds, particularly in retroflex or bunched articulations.
  - Duration helps in distinguishing rhotics from adjacent vowels. Rhotics, especially flaps, tend to have shorter durations.
  - To differentiate from adjacent vowels, a dip in intensity can help (yellow line in Praat picture).
  - It should be noted that a trill is much easier to differentiate from adjacent vowels than a flap because trills typically consist of multiple rapid closures of the tongue

against the alveolar ridge, creating distinct bursts of energy visible in both the waveform and spectrogram. In contrast, flaps are characterized by a single, brief contact of the tongue, resulting in a shorter and less acoustically distinct interruption, which makes them more challenging to segment from adjacent vowels.

- A radical drop in amplitude can be a cue to where rhotics begin, as vowels have a higher amplitude than rhotics.
- F4-F5 range: Subtle acoustic differences in F4 and F5 provide additional information when coding the *Surface Form* of rhotics, particularly for distinguishing bunched and retroflex variants (Zhou et al. 2008: 4467).

Lateral Laterals are approximants, which means they are among the most sonorous consonants. As their classification hints, they are somewhere between a vowel and a consonant, a fact that can make it hard to differentiate between the sounds. Figure B·2 includes the lateral /l/ in the word [pala].



(a) *Source:* SIB 0049PR

Figure B·2: Waveform, spectrogram, and annotation of the word [pala]

- The following cues can be used to find and segment the lateral in order to differentiate between the previous and following sounds:
  - The lateral's waveform is clearly periodic, a characteristic that can help differentiate it from flanking consonants. However, periodicity alone does not distinguish laterals from vowels, as both exhibit periodic waveforms. Additional acoustic cues, such as differences in intensity, formant transitions, and amplitude changes, are critical for segmentation.
  - For a lateral, there is no major dip in amplitude or intensity *(yellow line in Praat picture)* in contrast to rhotics, which tend to have lower intensity relative to vowels. This distinction is particularly relevant in contexts where liquids may exhibit overlapping acoustic properties, such as when a purportedly phonological rhotic sound is realized with lateral-like qualities.
  - When differentiating between a lateral and a vowel, the lateral typically shows slightly lower intensity due to the inherently higher sonority of vowels. This difference often appears as a subtle dip in intensity at the boundary between the two sounds. Additionally, while both vowels and laterals exhibit formant transitions, those of laterals tend to be more gradual and less pronounced.
  - In post-vocalic positions, laterals often exhibit abrupt shifts in formant structure, providing a key cue for segmentation in these contexts.

It should be noted that the goal of the study is to examine variation in spontaneous speech. As such, "idealized" tokens, like the ones shown in Figures B·1 and B·2, will rarely be encountered. The ultimate goal is to make the best segmentation possible, within reason.

Data Extraction and Cleaning

The data processing workflow consists of two main steps: extraction and cleaning. These steps ensure a systematic approach to preparing the data for analysis, minimizing human error and automating repetitive tasks.

Step 1: Data Extraction with Praat Script

The first step involves extracting annotated data from Praat TextGrids using a custom *Praat Script*, `Liquids_Data_Extraction.praat`. This script extracts all relevant tiers and acoustic measurements, compiling them into a comprehensive CSV file. Additionally, it converts coded annotations from the coding manual into human-readable labels and performs the following tasks:

- Extracts tier labels, acoustic measurements (e.g., formants, segment durations), and contextual information.
- Converts Spanish diacritics (e.g., accents) and IPA codes for seamless compatibility with downstream analysis in **R**.
- Handles multiple tokens within a single word and calculates additional metrics, such as segment and word durations.
- Generates a structured CSV file containing fields for speaker metadata, phonological forms, surface forms, contextual sounds, and acoustic measurements at 11 equidistant time points (F1–F5).

The extraction script reduces manual errors, streamlines the data preparation process,

and ensures consistency across all files. Detailed comments embedded in the script outline its functionality and modifications.

Step 2: Data Cleaning with R Script

Once the data has been extracted into a CSV file, the second step is data cleaning and transformation in **R**, using the script `Create_liquids_master_df.Rmd`. This script further refines the data and performs the following tasks:

DATA TIDYING - Removes unnecessary whitespace, fills undefined values with `NA`, and corrects inconsistent labels.

TOKEN CATEGORIZATION - Separates word classes (e.g., `verb`, `noun`), creates phonological and surface form types (e.g., `lateral`, `rhotic`), and revalues contextual sound categories (e.g., `vowel`, `consonant`).

FEATURE ENGINEERING - Adds derived variables such as:

- SPEECH RATE - Calculates milliseconds per syllable based on word duration and syllable count.
- LEXICAL FREQUENCY - Counts occurrences of each word in the dataset.
- PHONETIC-PHONOLOGICAL AGREEMENT - Indicates whether surface and phonological forms align.
- CONTEXTUAL FEATURES - Extracts manner and place of articulation for preceding and following sounds, as well as vowel height and displacement.

DATA INTEGRATION - Merges linguistic data with speaker metadata (e.g., sociolinguistic background) for richer analysis.

EXPORT - Saves the cleaned and aggregated data to a new CSV file for downstream analysis.

This two-step process ensures that the final dataset is both comprehensive and ready for variationist and acoustic analysis.

Bibliography

Alba, Orlando (1990). *Variación fonética y diversidad social en el español dominicano*. Spanish. URL: http://dialnet.unirioja.es/servlet/exttes?codigo=41966.

Alba, Orlando (2004). *Cómo hablamos los dominicanos: un enfoque sociolingüístico*. Santo Domingo: Grupo León Jimenes.

Babel, Anna M., ed. (2016a). *Awareness and Control in Sociolinguistic Research*. 1st edition. Cambridge University Press. DOI: 10.1017/CBO9781139680448.

Babel, Anna M. (2016b). "Silence as control: Shame and self-consciousness in sociolinguistic positioning". In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 9, pp. 200–227. DOI: 10.1017/CBO9781139680448.011.

Bailey, Karl G D & Ferreira, Fernanda (2003). "Disfluencies affect the parsing of garden-path sentences". In: *Journal of Memory and Language* 49.2, pp. 183–200.

Barrenechea, Ana María & Alonso, Alicia (1973). "Los pronombres personales sujetos en el español hablado en Buenos Aires". In: *Studia iberica: Festschrift für Hans Flasche*. Ed. by Juan M Lope Blanch. Mexico City: Universidad Nacional Autónoma de México: Bern, München, Francke, pp. 75–91.

Bartoń, Kamil (2023). *MuMIn: Multi-Model Inference*.

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve (2015). "Fitting Linear Mixed-Effects Models Using {lme4}". In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Bayley, Robert & Pease-Alvarez, Lucinda (1997). "Null pronoun variation in Mexican-descent children's narrative discourse". In: *Language Variation and Change* 9.3, pp. 349–371. DOI: 10.1017/S0954394500001964.

Beaton, Mary Elizabeth (2005). “Coda Liquid Production and Perception in Puerto Rican Spanish”. Ph.D. The Ohio State University. ISBN: 978-1-339-19201-7. URL: https://www.proquest.com/docview/1733330470.

Beaton, Mary Elizabeth (2016). “Revisiting Incomplete Neutralization: The Case of Puerto Rican Spanish”. In: *University of Pennsylvania Working Papers in Linguistics*. Vol. 22. 1, pp. 31–40. URL: https://repository.upenn.edu/pwpl/vol22/iss1/5.

Beck, Erica (2016). “What It Means to Be an Outsider: How Exposure to Regional Variation Shapes Children’s Awareness of Regional Accents in Their Native Language”. In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 5, pp. 104–122. DOI: 10.1017/CBO9781139680448.007.

Bell, Alan & Jurafsky, Daniel & Fosler-Lussier, Eric & Girand, Cynthia & Gregory, Michelle & Gildea, Daniel (2003). “Effects of disfluencies, predictability, and utterance position on word form variation in English conversation”. In: *The Journal of the Acoustical Society of America* 113.2. ISSN: 0001-4966.

Benevento, Nicole M & Dietrich, Amelia J (2015). “I think, therefore digo yo: Variable position of the 1sg subject pronoun in New Mexican Spanish-English code-switching”. In: *International Journal of Bilingualism* 19.4, pp. 407–422. ISSN: 1367-0069.

Bentivoglio, Paola (1987). *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela. ISBN: 980-00-0165-4.

Boersma, Paul & Weenink, David (2020). *Praat: doing phonetics by computer*. URL: https://www.praat.org.

Brown, Esther L (2004). “The reduction of syllable-initial /s/ in the Spanish of New Mexico and southern Colorado: A usage-based approach”. PhD thesis. The University of New Mexico.

Brown, Esther L & Torres Cacoullos, Rena (2003). “Spanish /s/: A different story from beginning (initial) to end (final)*”. In: *A Romance Perspective on Language Knowledge and Use: Selected papers from the 31st Linguistic Symposium on Romance Languages (LSRL), Chicago, 19–22 April 2001*. Ed. by Rafael Núñez-Cedeño & Luis López & Richard Cameron. John Benjamins Publishing Company, pp. 21–38. DOI: 10.1075/cilt.238.05bro.

Bucholtz, Mary & Hall, Kira (2005). “Identity and interaction: A sociocultural linguistic approach”. In: *Discourse Studies* 7.4-5, pp. 585–614. DOI: 10.1177/1461445605054407.

Bybee, Joan L (2002). “Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change”. In: *Language Variation and Change* 14.2, pp. 261–290. DOI: 10.1017/S0954394502143018.

Cameron, Richard (1992). “Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation”. PhD thesis. University of Pennsylvania.

Cameron, Richard (1993). “Ambiguous Agreement, Functional Compensation, and Nonspecific Tu in the Spanish of San Juan, Puerto Rico, and Madrid, Spain”. In: *Language Variation and Change* 5.3, pp. 305–334. ISSN: 14698021. DOI: 10.1017/S0954394500001526.

Cameron, Richard (1995). “The Scope and Limits of Switch Reference as a Constraint on Pronominal Subject Expression”. In: *Hispanic Linguistics* 6/7, pp. 1–28.

Campos-Astorkiza, Rebeka (2012). “The phonemes of Spanish”. In: *The Handbook of Hispanic Linguistics*. Ed. by José Ignacio Hualde & Antxon Olarrea & Erin O’Rourke. Wiley-Blackwell. Chap. 5, pp. 89–110.

Carmichael, Katie (2016). “Place-linked expectations and listener awareness of regional accents”. In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 7, pp. 152–176. DOI: 10.1017/CBO9781139680448.009.

Carvalho, Ana M & Child, Michael (2011). “Subject Pronoun Expression in a Variety of Spanish in Contact with Portuguese”. In: *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Ed. by Jim Michnowicz & Robin Dodsworth. Somerville, MA: Cascadilla Proceedings Project, pp. 14–25.

Cerrón-Palomino, Álvaro (2014). “Ser más PRO o menos PRO: Variación en la expresión de sujeto pronominal en el castellano limeño”. In: *Lingüística* 30.1, pp. 61–83.

Choksi, Nishaant & Meek, Barbra A. (2016). “Theorizing salience: Orthographic practice and the enfigurement of minority languages”. In: *Awareness and Control in Sociolinguistic*

Research. Ed. by Anna M. Babel. Cambridge University Press. Chap. 10, pp. 228–252. DOI: 10.1017/CBO9781139680448.012.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. MIT Press.

Clark, Herbert H & Fox Tree, Jean E (2002). "Using uh and um in spontaneous speaking". In: *Cognition* 84.1, pp. 73–111. ISSN: 0010-0277.

Cole, Amanda (2020). "Co-variation, style and social meaning: The implicational relationship between (h) and (ing) in Debden, Essex". In: *Language Variation and Change* 32.3, pp. 349–371. DOI: 10.1017/S0954394520000162.

Couper-Kuhlen, Elizabeth & Ono, Tsuyoshi (2007). "'Incrementing' in conversation. A comparison of practices in English, German and Japanese". In: *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)* 17.4, pp. 513–552.

de Prada Pérez, Ana (2009). "Subject expression in Minorcan Spanish: Consequences of contact with Catalan". PhD thesis. ISBN: 9781109385601. URL: http://search.proquest.com/docview/304981820/.

de Prada Pérez, Ana (2020). "The interaction of functional predictors and the mechanical predictor perseveration in a variationist analysis of caribbean spanish heritage speaker subject pronoun expression". In: *Languages* 5.4, pp. 1–20. ISSN: 2226471X. DOI: 10.3390/languages5040036.

Delgado-Díaz, Gibran & Galarza, Iraida (2015). "¿Que comiste [x]amón? A Closer Look at the Neutralization of /h/ and Posterior /r/ in Puerto Rican Spanish". In: 2013, pp. 70–82.

Dionne, Danielle (2023). "Contextual Frequency and Morphosyntactic Variation: An exemplar-theoretic variationist analysis of Spanish Subject Pronouns". PhD thesis. Boston University, pp. 1–186.

Eckert, Penelope (1988). "Adolescent social structure and the spread of linguistic change". In: *Language in Society* 17.2, pp. 183–207. DOI: 10.1017/S0047404500012756.

Eckert, Penelope (2008). "Variation and the indexical field". In: *Journal of Sociolinguistics* 12.4, pp. 453–476. DOI: 10.1111/j.1467-9841.2008.00374.x.

Erker, Daniel (2010). “A subsegmental approach to coda /s/ weakening in Dominican Spanish”. In: *International Journal of the Sociology of Language* 203, pp. 9–26. DOI: 10.1515/IJSL.2010.019.

Erker, Daniel (2012). “An Acoustically Based Sociolinguistic Analysis of Variable Coda /s/ Production in the Spanish of New York City”. English. PhD thesis. New York University, p. 485.

Erker, Daniel (2017). “Contact, Co-Variation, and Sociolinguistic Salience: What Mister Rogers Knows about Language Change”. In: *University of Pennsylvania Working Papers in Linguistics (Selected Papers from NWAV45)* 23.2, pp. 1–13. URL: https://repository.upenn.edu/pwpl/vol23/iss2/9.

Erker, Daniel (2022). “How social salience can illuminate the outcomes of linguistic contact: Data from Spanish in Boston”. In: *The Coherence of Linguistic Communities: Orderly Heterogeneity and Social Meaning*. Ed. by Karen V. Beaman & Gregory R. Guy. 1st. New York: Routledge. Chap. 8, pp. 145–162. DOI: 10.4324/9781003134558-12.

Erker, Daniel & Bruso, Joanna (2017). “Uh, bueno, em: Filled pauses as a site of contact-induced change in Boston Spanish”. In: *Language Variation and Change* 29.2, pp. 205–244. ISSN: 14698021. DOI: 10.1017/S0954394517000102.

Erker, Daniel & Ho-Fernández, Eduardo & Otheguy, Ricardo & Shin, Naomi (2017). “Continuity and change in Spanish among Cubans in New York: A study of subject placement with finite verbs”. In: *Cuban Spanish Dialectology: Variation, Contact, and Change*. Ed. by Alejandro Cuza. Washington, DC: Georgetown University Press. Chap. 4, pp. 61–80. ISBN: 9781626165113.

Erker, Daniel & Guy, Gregory R (2012). “The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish”. In: *Language* 88.3, pp. 526–557. DOI: 10.1353/lan.2012.0050.

Erker, Daniel & Otheguy, Ricardo (2016). “Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish”. English. In: *Lingua* 172-173, pp. 131–146. DOI: 10.1016/j.lingua.2015.10.011.

Erker, Daniel & Otheguy, Ricardo (2020). “American myths of linguistic assimilation: A sociolinguistic rebuttal”. In: *Language in Society*, pp. 1–37.

Erker, Daniel & Reffel, Madeline (2021). “Describing and analyzing variability in Spanish /s/: A case study of Caribbeans in Boston and New York City”. In: *Sociolinguistic Approaches to Sibilant Variation in Spanish*. Ed. by Eva Núñez-Méndez. 1st. New York: Routledge. Chap. 4, pp. 131–163. DOI: 10.4324/9781003153948.

Erker, Daniel & Vidal-Covas, Lee-Ann (2022). “What We Say When We Say Nothing at All: Clues to Contact-Induced Language Change in Spanish Conversational Pause-Fillers”. In: *Estudios del Observatorio/Observatorio Studies* 80, pp. 1–29. DOI: 10.15427/OR080-09/2022EN.

Erker, Daniel & Vidal-Covas, Lee-Ann (2024). “Variation, contact, and change in Boston Spanish: how social meaning shapes stylistic practice and bilingual optimization”. In: *Studies in Hispanic and Lusophone Linguistics* 17.2, pp. 223–249. DOI: 10.1515/shll-2024-2010.

Escobar, Anna María (2012). “Spanish in contact with Amerindian languages”. In: *The Handbook of Hispanic Linguistics*. Ed. by José Ignacio Hualde & Antxon Olarrea & Erin O’Rourke. Wiley-Blackwell. Chap. 4, pp. 65–88.

Escobar, Anna María & Potowski, Kim (2015). *El español de los Estados Unidos*. Cambridge University Press.

Flores-Ferrán, Nydia (2002). *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective*. Lincom Europa.

Flores-Ferrán, Nydia (2004). “Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact?” In: *Language Variation and Change* 16.1, pp. 47–73. DOI: 10.1017/S0954394504161048.

Flórez, Luis (1951). *La pronunciación del español en Bogotá*. Vol. 8. Bogotá: Instituto Caro y Cuervo.

Fox, Jon & Weisberg, Sanford (2019). *An {R} Companion to Applied Regression*. Third. Sage Publications. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Fruehwald, Josef (2016). “Filled Pause Choice as a Sociolinguistic Variable”. In: *University of Pennsylvania Working Papers in Linguistics* 22.2, pp. 41–49. URL: https://repository.upenn.edu/pwpl/vol22/iss2/6.

Galué, Dexy (2002). “Marcadores conversacionales: un análisis pragmático”. In: *Boletín de Lingüística* 18, pp. 27–48.

Gilbert, Madeline B & Rohena-Madrazo, Marcos (2017). “Revising the canon Social and stylistic variation of coda /ɾ/ in Buenos Aires Spanish”. In: *Romance Languages and Linguistic Theory 12: Selected papers from the 45th Linguistic Symposium on Romance Languages (LSRL)* 12. Ed. by Ruth E.V. University of Campinas) Lopes & Juanito (University of Campinas) Ornelas de Avelar & Sonia M. L. (University of Campinas) Cyrino, pp. 63–78. ISSN: 24056944. DOI: 10.1075/rllt.12.05gil.

Gradoville, Michael S. & Brown, Earl Kjar & File-Muriel, Richard J. (2022). “The phonetics of sociophonetics: Validating acoustic approaches to Spanish /s/”. In: *Journal of Phonetics* 91, pp. 1–32. DOI: 10.1016/j.wocn.2021.101125.

Graham, Lamar (2013). “Comparing hesitation markers in Sanjuanero Spanish”. In: *Diálogo de la Lengua* 5, pp. 66–77.

Granda, Germán de (1991). *El español en tres mundos: retenciones y contactos lingüisticos en América y África*. Valladolid: Universidad de Valladolid, Secretariado de Publicaciones.

Guy, Gregory R (2013). “The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables?” In: *Journal of Pragmatics* 52, pp. 63–71. DOI: 10.1016/j.pragma.2012.12.019.

Guy, Gregory R & Hinskens, Frans (2016). “Linguistic coherence: Systems, repertoires and speech communities”. In: *Lingua* 172-173, pp. 1–9. DOI: 10.1016/j.lingua.2016.01.001.

Guy, Gregory R & Oushiro, Livia & Mendes, Ronald Beline (2022). “Indexicality and coherence”. In: *The Coherence of Linguistic Communities: Orderly Heterogeneity and Social Meaning*. Ed. by Karen V. Beaman & Gregory R. Guy. 1st. New York: Routledge. Chap. 3, pp. 53–68. DOI: 10.4324/9781003134558-5.

Hawkins, John A & Filipović, Luna (2024). “Bilingualism-induced language change: what can change, when, and why?” In: *Linguistics Vanguard* 10.s2, pp. 115–124. ISSN: 2199174X. DOI: 10.1515/lingvan-2023-0089.

Hualde, José Ignacio (2014). *Los sonidos del español*. Cambridge University Press, p. 361. ISBN: 9781107595446. DOI: 10.1017/CBO9780511719943.

Katz, Jonah (2021). “Intervocalic lenition is not phonological: Evidence from Campidanese Sardinian”. In: *Phonology* 38.4, pp. 651–692. ISSN: 14698188. DOI: 10.1017/S095267572100035X.

Labov, William (1963). “The Social Motivation of a Sound Change”. In: *WORD* 19.3, pp. 273–309. DOI: 10.1080/00437956.1963.11659799.

Labov, William (1964). “The Social Stratification Of English In New York City”. PhD thesis. New York: Columbia University. URL: https://www.proquest.com/pqdtglobal/docview/302149177/90B434AA9C2E4D48PQ/12?accountid=9676.

Labov, William (1966). *The Social Stratification of English in New York City*. 2nd edition. Cambridge University Press. ISBN: 0521821223. DOI: 10.1017/CBO9780511618208.

Labov, William (1972). *Sociolinguistic Patterns*. 4. University of Pennsylvania Press. ISBN: 9780812210521.

Labov, William (2001). *Principles of Linguistic Change: Social Factors*. Vol. 2. Blackwell Publishing. ISBN: 0631179151.

Lawson, Eleanor & Stuart-Smith, Jane & Scobbie, James M. & Yaeger-Dror, Malcah & Maclagan, Margaret (2010). “Analyzing Liquids”. In: *Sociophonetics: A Student’s Guide*. Ed. by Marianna Di Paolo & Malcah Yaeger-Dror. Routledge, pp. 72–86. ISBN: 9780415498784.

Lipski, John M (1985). *The Spanish of Equatorial Guinea: the dialect of Malabo and its implications for Spanish dialectology*. Tübingen, Germany.

Lipski, John M (1994). “Tracing Mexican Spanish /s/:: A Cross-Section of History Tracing Mexican Spanish /s/:: A Cross-Section of History Tracing Mexican Spanish /s/:: A Cross-Section of History”. In: *Language Problems and Language Planning* 18.3, pp. 223–241.

Lipski, John M (1996a). “Contactos de criollos en el Caribe hispánico: contribuciones al español bozal”. In: *América negra* 11, pp. 31–60.

Lipski, John M (1996b). *El español de América*. Madrid, España: Ediciones Cátedra.

Lipski, John M (1999). “The many faces of Spanish /s/-weakening: (Re) alignment and ambisyllabicity”. In: *Advances in Hispanic linguistics*, pp. 198–213.

Lipski, John M (2001). “The place of Chabacano in the Philippine linguistic profile”. In: *Sociolinguistic Studies* 2.2, pp. 119–163.

Lipski, John M (2003). “La lengua española en los Estados Unidos: avanza a la vez que retrocede”. In: *Revista española de lingüística* 33.2, pp. 231–260.

Lipski, John M (2005a). “Code-switching or borrowing? No sé so no puedo decir, you know”. In: *Selected proceedings of the second workshop on Spanish sociolinguistics*. Cascadilla Proceedings Project, pp. 1–15.

Lipski, John M (2005b). “El español en el mundo: frutos del último siglo de contactos lingüisticos”. In: *Contactos y contextos lingüísticos: el español en los Estados Unidos y en contacto con otras lenguas*, pp. 29–53.

Lipski, John M (2008). *Varieties of Spanish in the United States*. Georgetown Studies in Spanish Linguistics. Washington, D.C.: Georgetown University Press. ISBN: 9781589012134.

Lipski, John M (2012). “Geographical and social varieties of Spanish: An overview”. In: *The Handbook of Hispanic Linguistics*. Ed. by José Ignacio Hualde & Antxon Olarrea & Erin O’Rourke. Wiley-Blackwell. Chap. 1, pp. 1–26.

López Morales, Humberto (1983). “La lateralización de la /r/ en el español de Puerto Rico”. In: *Philologica hispaniensia : in honorem Manuel Alvar*. Madrid: Gredos, pp. 387–398. ISBN: 9788424908997.

López Morales, Humberto (2009). “Introducción: presencia histórica de lo hispano”. In: *Enciclopedia del español en los Estados Unidos*. Ed. by Humberto López Morales. Madrid, España: Santillana. Chap. Introducci, p. 31.

Maclay, Howard & Osgood, Charles E (1959). “Hesitation Phenomena in Spontaneous English Speech”. In: *WORD* 15.1, pp. 19–44. ISSN: 0043-7956.

Martínez-Gil, Fernando (2012). “Main Phonological Processes”. In: *The Handbook of Hispanic Linguistics*. Ed. by José Ignacio Hualde & Antxon Olarrea & Erin O’Rourke. Wiley-Blackwell, pp. 111–132.

McGowan, Kevin B. (2016). “Sounding Chinese and listening Chinese: Awareness and knowledge in the laboratory”. In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 2, pp. 25–61. DOI: 10.1017/CBO9781139680448.004.

Mendoza-Denton, Norma (2008). *Homegirls: Language and Cultural Practice Among Latina Youth Gangs*. Blackwell Publishing. DOI: 10.1002/9780470693728.

Michnowicz, Jim (2015). “Subject Pronoun Expression in Contact with Maya in Yucatan Spanish”. In: *Subject pronoun expression in Spanish : a cross-dialectal perspective*. Ed. by Naomi Ladipus Shin & Rafael Orozco & Ana Maria Carvalho. Georgetown studies in Spanish linguistics. Washington, DC: Georgetown University Press, pp. 101–120. ISBN: 9781626161719.

Morales, Amparo (1982). “La perspectiva dinámica oracional en el español de Puerto Rico”. In: *El español del Caribe (Ponencias del VI Simposio de Dialectología)*. Ed. by Orlando Alba. Santiago de los Caballeros, República Dominicana: Universidad Católica Madre y Maestra, pp. 203–219.

Morales, Amparo (1999). “Anteposición del sujeto en el español del Caribe”. In: *El Caribe hispánico: Perspectivas lingüísticas actuales*, pp. 77–98.

Morales de Walters, Amparo (1997). “La hipótesis funcional y la aparición de sujeto no nominal: el español de Puerto Rico”. In: *Hispania* 80.1, pp. 153–165. DOI: 10.2307/345995.

Muysken, Pieter. (2013). “Language contact outcomes as the result of bilingual optimization strategies”. In: *Bilingualism: Language and Cognition 16* 16.4, pp. 709–730. DOI: 10.1017/S1366728912000727.

Navarro Tomás, Tomás (1948). *El español en Puerto Rico. Contribución a la geografía lingüística hispanoamericana*. Universidad de Puerto Rico.

Nycz, Jennifer (2016). “Awareness and acquisition of new dialect features”. In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 3, pp. 62–79. DOI: 10.1017/CBO9781139680448.005.

Orozco, Rafael (2015). “Pronominal variation in Colombian costeño Spanish”. In: *Subject pronoun expression in Spanish: A cross-dialectal perspective*, pp. 17–37.

Orozco, Rafael & Hurtado, Luz Marcela (2021a). “A variationist study of subject pronoun expression in Medellín, Colombia”. In: *Languages* 6.1, pp. 1–29. ISSN: 2226471X. DOI: 10.3390/languages6010005.

Orozco, Rafael & Hurtado, Luz Marcela (2021b). “Variable subject pronoun expression revisited: This is what the Paisas do”. In: *Proceedings of the Linguistic Society of America (PLSA)*. Vol. 6. 1, pp. 713–721.

Orozco, Rafael & Mendez-Vallejo, Catalina & Vidal-Covas, Lee-Ann (2014). “Los efectos condicionantes del verbo en el uso variable de los pronombres personales de sujeto”. In: *Actas del XVII Congreso internacional de la Asociación de Lingüística y Filología de America Latina (ALFAL)*. João Pessoa - Paraíba, Brasil, pp. 1–17.

Ortiz López, Luis A (2009). “El español del Caribe: orden de palabras a la luz de la interfaz léxico-sintáctica y sintáctico-pragmática”. In: *Revista Internacional de Lingüística Iberoamericana* 7.2 (14), pp. 75–93. URL: https://www.jstor.org/stable/41678402.

Ortiz López, Luis A (2019). “Pronombres de sujeto en el español (L2 vs. L1) del Caribe”. In: *Español en Estados Unidos y otros contextos de contacto*, pp. 85–110. DOI: 10.31819/9783865279033-006.

Otheguy, Ricardo & Zentella, Ana Celia (2012). *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199737406.001.0001.

Otheguy, Ricardo & Zentella, Ana Celia & Livert, David (2007). “Language and Dialect Contact in Spanish in New York: Toward the Formation of a Speech Community”. In: *Language* 83.4, pp. 770–802. URL: http://www.jstor.org/stable/40070965.

Oushiro, Livia (2016). “Social and structural constraints in lectal cohesion”. In: *Lingua* 172-173, pp. 116–130. DOI: 10.1016/j.lingua.2015.10.015.

Oushiro, Livia & Guy, Gregory R (2015). “The Effect of Salience on Co-variation in Brazilian Portuguese”. In: *University of Pennsylvania Working Papers in Linguistics* 21.2, pp. 157–166. URL: https://repository.upenn.edu/handle/20.500.14332/45063.

Penny, Ralph (2000). *Variation and Change in Spanish*. Cambridge University Press. ISBN: 0521780454.

Perlmutter, David & Postal, Paul (1974). “Lectures on relational grammar”. unpublished notes, Summer Linguistics Institute of the LSA. Amherst, MA.

Pluymaekers, Mark & Ernestus, Mirjam & Baayen, R Harald (2005). “Lexical frequency and acoustic reduction in spoken Dutch”. In: *The Journal of the Acoustical Society of America* 118.4, pp. 2561–2569. DOI: 10.1121/1.2011150.

Poplack, Shana (1979). “Function and process in a variable phonology”. PhD thesis. University of Pennsylvania.

Poplack, Shana (2001). “Variability, frequency, and productivity in the irrealis domain of French”. In: *Frequency and the Emergence of Linguistic Structure*. Ed. by Joan L Bybee & Paul J Hopper. John Benjamins Publishing Company, pp. 405–428. DOI: doi.org/10.1075/tsl.45.20pop.

Poplack, Shana & Levey, Stephen (2010). “Contact-induced grammatical change: A cautionary tale”. In: *Language and space: An international handbook of linguistic variation*. Ed. by P Auer & J E Schmidt. Vol. 1. Theories and methods. Mouton de Gruyte, pp. 391–419.

Posio, Pekka (2014). “Subject expression in grammaticalizing constructions: The case of creo and acho ‘I think’ in Spanish and Portuguese”. In: *Journal of Pragmatics* 63, pp. 5–18. ISSN: 03782166. DOI: 10.1016/j.pragma.2013.07.001.

Preston, Dennis R (2016). “Whaddayaknow now?” In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 8, pp. 177–199. DOI: 10.1017/CBO9781139680448.010.

Prosper-Sánchez, Gloria D (1995). “Neutralización Homofonética de líquidas a final de silaba: Aspectos sociolingüísticos en el español de Puerto Rico”. Dissertation. University of Massachusetts Amherst. ISBN: 0315880678. DOI: 10.1016/j.jaci.2012.05.050.

Quilis, Antonio (1988). “Nuevos datos sobre la actitud de los ecuatoguineanos ante la lengua española”. In: *Nueva Revista de Filologia Hispánica* 36.2, pp. 719–731. ISSN: 01850121. URL: http://www.jstor.org/stable/40300285.

Quilis, Antonio (1992). *La lengua española en cuatro mundos*. 4. Madrid: Editorial MAPFRE. ISBN: 9788471005229.

Quilis, Antonio (1993). *Tratado de fonética y fonología españolas*. Gredos. ISBN: 978-8424939212.

R, Core Team (2024). *R: A language and environment for statistical computing*. Vienna, Austria. URL: https://www.r-project.org/.

Rácz, Péter (2013). *Salience in Sociolinguistics: A Quantitative Approach*. De Gruyter Mouton. DOI: 10.1515/9783110305395.

Ramos-Pellicia, Michelle F (2004). "Language Contact and Dialect Contact: Cross-Generational Phonological Variation in a Puerto Rican Community in the Midwest of the United States". PhD thesis. The Ohio State University. URL: https://www.proquest.com/docview/85634668.

Ramos-Pellicia, Michelle F (2007). "Lorain Puerto Rican Spanish and "r" in three generations". In: *Selected Proceedings of the 3rd Workshop on Spanish Sociolinguistics*. Ed. by Jonathan Holmquist & Augusto Lorenzino & Lotfi Sayahi. January 2007. Cascadilla Proceedings Project, pp. 53–60. URL: http://www.lingref.com/cpp/wss/3/paper1526.pdf.

Raña-Risso, Rocío (2013). "A Corpus-Based Sociolinguistic Study of Subject Pronoun Placement in Spanish in New York". PhD thesis. City University of New York ProQuest Dissertations Publishing.

Raña-Risso, Rocío & Barrera-Tobón, Carolina (2018). "On the relationship between subject placement and overt pronouns in the Spanish of New York City bilinguals". In: *Journal of Language Contact* 11.2, pp. 324–347. DOI: 10.1163/19552629-01102007.

Sayahi, Lofti (2004). "The Spanish language presence in Tangier, Morocco: a sociolinguistic perspective". In: *Afro-Hispanic Review* 23.2, pp. 54–61. ISSN: 02788969. URL: http://www.jstor.org/stable/23054554.

Schilling, Natalie (2013). "Investigating Stylistic Variation". In: *The Handbook of Language Variation and Change*. Ed. by J. K. Chambers & Natalie Schilling. 2nd edition. John Wiley & Sons. Chap. 15, pp. 325–349. DOI: 10.1002/9781118335598.ch15.

Sharma, Devyani (2022). "False oppositions in the study of coherence". In: *The Coherence of Linguistic Communities: Orderly Heterogeneity and Social Meaning*. Ed. by Karen V.

Beaman & Gregory R. Guy. 1st. New York: Routledge. Chap. 1, pp. 17–33. DOI: 10.4324/9781003134558-3.

Shenk, Elaine (2021). “Eso es puertorriqueño: The enregisterment of lateralization among puerto ricans in the United States”. In: *Centro Journal* 33.2, pp. 126–165. ISSN: 21632960.

Shin, Naomi Ladipus & Orozco, Rafael & Carvalho, Ana Maria, eds. (2015). *Subject Pronoun Expression in Spanish: A Cross-dialectal Perspective*. Washington, DC: Georgetown University Press. ISBN: 9781626161719.

Shin, Naomi Lapidus & Erker, Daniel (2015). “The emergence of structured variability in morphosyntax: Childhood acquisition of Spanish subject pronouns”. In: *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Ed. by Ana M. Carvalho & Rafael Orozco & Naomi Lapidus Shin, pp. 169–189.

Shin, Naomi Lapidus & Montes-Alcalá, Cecilia (2014). “El uso contextual del pronombre sujeto como factor predictivo de la influencia del inglés en el español de nueva york”. In: *Sociolinguistic Studies* 8.1, pp. 85–110. ISSN: 17508657. DOI: 10.1558/sols.v8i1.85.

Silva Corvalán, Carmen (1982). “Subject expression and placement in Mexican-American Spanish”. In: *Spanish in the United States: Sociolinguistic Aspects*. Ed. by Jon Amastae & Lucia Elías-Olivares. Cambridge University Press. Chap. 5, pp. 93–120.

Silva-Corvalán, Carmen (1994). *Language contact and change: Spanish in Los Angeles*. New York, NY: Oxford University Press. ISBN: 0198236441.

Silva-Corvalán, Carmen (2008). *The Limits of Convergence in Language Contact*. Vol. 2. 1. Leiden, The Netherlands: Brill, pp. 213–224.

Silverstein, Michael (2001). “The limits of awareness (reprint from 1981)”. In: *Linguistic Anthropology: A Reader*. Ed. by Alessandro Duranti. Vol. 1. Blackwell Publishers Ltd, pp. 382–401. DOI: 10.1016/s0262-4079(23)01756-6.

Solon, Megan (2015). “L2 Spanish /l/: The Roles of F2 and Segmental Duration in Foreign Accent Perception”. In: *Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology*. Ed. by Erik W. Willis & Pedro Martín Butragueño & Esther Herrera Zendejas. Cascadilla Proceedings Project, pp. 83–94. URL: http://www.lingref.com/cpp/larp/6/paper3194.pdf.

Sorenson, Travis. (2013). “Voseo to Tuteo accommodation among Salvadorans in the United States”. In: *Hispania* 96.4, pp. 763–781.

Squires, Lauren (2016). “Processing Grammatical Differences: Perceiving versus Noticing”. In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 4, pp. 80–103. DOI: 10.1017/CBO9781139680448.006.

Swerts, Marc (1998). “Filled pauses as markers of discourse structure”. In: *Journal of Pragmatics* 30.4, pp. 485–496. ISSN: 0378-2166.

Tamminga, Meredith & Wade, Lacey (2022). “Coherence across social and temporal scales”. In: *The Coherence of Linguistic Communities: Orderly Heterogeneity and Social Meaning*. Ed. by Karen V. Beaman & Gregory R. Guy. 1st. New York: Routledge. Chap. 2, pp. 34–52. DOI: 10.4324/9781003134558-4.

Toribio, Almeida Jaqueline (2004). “Convergence as an optimization strategy in bilingual speech: Evidence from code-switching”. In: *Bilingualism: Language and Cognition* 7.2, pp. 165–173. DOI: 10.1017/S1366728904001476.

Torres Cacoullos, Rena & Travis, Catherine E (2010). “Testing convergence via code-switching: priming and the structure of variable subject expression”. In: *International Journal of Bilingualism* 15.3, pp. 241–267. ISSN: 1367-0069.

Torres Cacoullos, Rena & Travis, Catherine E (2016). “Two languages, one effect: Structural priming in spontaneous code-switching”. In: *Bilingualism* 19.4, pp. 733–753. ISSN: 13667289. URL: http://search.proquest.com/docview/1857295484/.

Torres Cacoullos, Rena & Travis, Catherine E (2018). *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge University Press. DOI: 10.1017/9781108235259.

Torres Cacoullos, Rena & Travis, Catherine E (2019). “Variationist typology: Shared probabilistic constraints across (non-) null subject languages”. In: *Linguistics* 57.3, pp. 653–692.

Tottie, Gunnel (2014). “On the use of uh and um in American English”. In: *Functions of Language* 21.1, pp. 6–29. ISSN: 0929-998X. URL: http://search.proquest.com/docview/1558997884/.

Trudgill, Peter (1986). *Dialects in contact*. Language in society ; 10. London: Wiley-Blackwell. ISBN: 978-0631127338.

U.S. Census Bureau (2019). *Facts on Latinos of Puerto Rican origin in the U.S.* Washington, D.C. URL: https://www.pewresearch.org/hispanic/fact-sheet/u-s-hispanics-facts-on-puerto-rican-origin-latinos/ (visited on 04/11/2022).

Valentín-Márquez, Wilfredo (2007). "Doing being boricua: Perceptions of national identity and the sociolinguistic distribution of liquid variables in Puerto Rican Spanish". PhD thesis. University of Michigan.

Valentín-Márquez, Wilfredo (2019). "The sociolinguistic distribution of Puerto Rican Spanish /r/ in Grand Rapids, Michigan". In: *Dialects from Tropical Islands*, pp. 88–110. DOI: 10.4324/9781315115443-6.

Vidal-Covas, Lee-Ann (2013). "El uso variable de los pronombres sujetos en el castellano puertorriqueño hablado en Luisiana y Puerto Rico". Spanish. PhD thesis. LSU. DOI: 10.31390/gradschool_theses.3876.

Vidal-Covas, Lee-Ann (2020). "Hesitation phenomena as a site of dialectal and language contact among Spanish-speaking Bostonians". Unpublished Master's Thesis. Boston University.

Vidal-Covas, Lee-Ann (2023). "¿Va primero el verbo or ¿El sujeto va primero?: Subject-verb order in Latin American Spanish". In: *Proceedings of the Linguistic Society of America (PLSA)* 8.1, pp. 1–17. DOI: 10.3765/plsa.v8i1.5542.

Watanabe, Michiko & Den, Yasuharu & Hirose, Keikichi & Minematsu, Nobuaki & Den, Yasuharu & Minematsu, Nobuaki (2004). "Clause types and filed pauses in Japanese spontaneous monologues". In: *INTERSPEECH*. Jeju Island, Korea.

Watanabe, Michiko & Hirose, Keikichi & Den, Yasuharu & Minematsu, Nobuaki (2008). "Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners". In: *Speech Communication* 50.2, pp. 81–94. ISSN: 0167-6393.

Willis, Erik W & Bradley, Travis G (2008). "Contrast maintenance of taps and trills in Dominican Spanish: Data and analysis". In: *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*. Ed. by Laura Colantoni & Jeffrey Steele.

Cascadilla Proceedings Project, pp. 87–100. URL: http://www.lingref.com/cpp/lasp/3/paper1716.pdf.

Woods, Michael R. & Rivera-Mills, Susana V. (2012). “El tú como un mask: Voseo and Salvadoran and Honduran Identity in the United States”. In: *Studies in Hispanic and Lusophone Linguistics* 5.1, pp. 191–216.

Zagona, Karen (2002). *The syntax of Spanish*. Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9780511613234.

Zapata, Gabriela & Sánchez, Liliana & Toribio, Almeida Jacqueline (2005). “Contact and contracting Spanish”. In: *International Journal of Bilingualism* 9.3–4, pp. 377–395. DOI: 10.1177/13670069050090030501.

Zhou, Xinhui & Espy-Wilson, Carol Y & Boyce, Suzanne & Tiede, Mark & Holland, Christy & Choe, Ann (2008). “A magnetic resonance imaging-based articulatory and acoustic study of “retroflex” and “bunched” American English /r/.” eng. In: *The Journal of the Acoustical Society of America* 123.6, pp. 4466–4481. DOI: 10.1121/1.2902168.

Zimman, Lal (2016). “Sociolinguistic agency and the gendered voice: Metalinguistic negotiations of vocal masculinization among female-to-male transgender speakers”. In: *Awareness and Control in Sociolinguistic Research*. Ed. by Anna M. Babel. Cambridge University Press. Chap. 11, pp. 253–277. DOI: 10.1017/CBO9781139680448.013.